

Sistemi Operativi T – Ingegneria Informatica

Corso tenuto da Anna Ciampolini

APPUNTI DI TEORIA PER L'ESAME SCRITTO

Appunti di Andrei Daniel Ivan, and Reidaniel.ivan@studio.unibo.it

per l'anno accademico 2018/2019

(nel caso troviate errori grammaticali, errori concettuali, cose poco chiare o aggiornamento delle slide, mandatemi una mail e aggiornerò il file)

Indice

1. Generalità di un sistema operativo	4
1.1 Cos'è un sistema operativo?	4
1.2 Evoluzione dei sistemi operativi	4
1.3 Interruzioni	4
1.4 I/O	4
1.5 Protezione	4
1.6 Strutture	5
1.7 Macchine virtuali	5
2. Processi e thread	6
2.1 Processi	6
2.2 Scheduling	6
2.3 Operazioni sui processi	6
2.4 Thread (o processi leggeri)	6
2.5 Interazione tra processi	7
2.6 Processi in UNIX	7
3. Scheduling della CPU	8
3.1 Generalità	8
3.2 Criteri di scheduling	8
3.3 Algoritmi di scheduling	8
4. File System	9
4.1 Generalità	9
4.2 File	9
4.2.1 Operazioni sui file	9
4.2.2 Struttura interna dei file	10
4.2.3 Metodi di accesso ai file	10
4.3 Directory	10
4.4 Organizzazione fisica del file system	10
4.5 File system in UNIX	11
4.5.1 Accesso ai file in UNIX	11
5. Gestione della memoria centrale	12
5.1 Generalità	12
5.2 Tecniche di allocazione in memoria centrale	12
5.2.1 Allocazione contigua	12
5.2.2 Allocazione non contigua	12
5.3 Memoria Virtuale	13
5.3.1 Paginazione su richiesta	13
5.3.2 Pre-paginazione e working set	14
5.4 Gestione della memoria in UNIX	14

6. Programmazione concorrente ad ambiente globale	15
6.1 Generalità	15
6.2 Deadlock	15
6.3 Mutua esclusione	15
7. Thread in Java	16
7.1 Generalità	16
7.2 Semafori	16
7.3 Monitor	16
8. Gestione I/O	17
8.1 Compiti del gestore di I/O	17
8.2 Architettura del sottosistema di I/O	17
8.3 Gestione e organizzazione fisica dei dischi	17

1. Generalità di un sistema operativo

1.1 Cos'è un sistema operativo?

Un sistema operativo è un programma, o un insieme di programmi, che fornisce all'utente una visione astratta e semplificata dell'hardware e gestisce in modo efficace ed efficiente le risorse del sistema. È quindi l'intermediario tra utente e hardware.

1.2 Evoluzione dei sistemi operativi

- **Sistemi Batch** (seconda generazione): in questo tipo di sistemi l'unico compito del sistema operativo è quello di trasferire il controllo da un job, appena terminato, al prossimo da eseguire del lotto, detto *batch*. La CPU è spesso inattiva a causa dell'attesa di eventi da parte dei job.
- **Sistemi multi-programmati** (terza generazione): viene precaricato sul disco un insieme di job, detto *pool*; il sistema operativo ha il compito di caricare un sottoinsieme di job appartenenti al pool e selezionare un job tra essi che verrà assegnato alla CPU. Qualora il job assegnato alla CPU si ponesse in attesa di un evento, il sistema operativo assegnerebbe la CPU ad un altro job. L'obiettivo della multiprogrammazione è la massimizzazione dell'utilizzo della CPU e della memoria.
- **Sistemi time-sharing** (quarta generazione): implementano le seguenti caratteristiche:
 - **multi-utenza**: il sistema operativo presenta ad ogni utente una macchina virtuale completamente dedicata;
 - **interattività**: Il sistema operativo interrompe l'esecuzione di un job dopo un intervallo di tempo prefissato, assegnando la CPU ad un altro job.

1.3 Interruzioni

Le varie componenti hardware e software interagiscono con il sistema operativo attraverso interruzioni asincrone (interrupt). Ogni interruzione è causata da un evento ed è associata ad un handler per la gestione di tale evento. Quando un programma chiede l'esecuzione di un servizio del sistema operativo parliamo di *system call*. Alla ricezione di una interruzione il sistema operativo interrompe la sua esecuzione salvandone lo stato in memoria, attiva successivamente l'handler specifico dell'evento e infine ripristina lo stato salvato in memoria.

1.4 I/O

L'I/O in un sistema di elaborazione avviene tramite dei controller, ovvero interfacce hardware delle periferiche verso il bus di sistema. Ogni controller è dotato di un registro dati, dove memorizzare le informazioni da leggere e scrivere, e di alcuni registri speciali, ove memorizzare le operazioni da eseguire (registro di controllo) e l'esito delle operazioni eseguite (registro di stato).

Il driver di dispositivo è un componente del sistema operativo che interagisce direttamente con il dispositivo e la sua struttura è strettamente dipendente da esso.

1.5 Protezione

Nei sistemi operativi che prevedono multiprogrammazione e multi-utenza sono necessari alcuni meccanismi di protezione. È necessario impedire al programma in esecuzione di accedere ad aree di memoria esterne al proprio spazio. Per garantire protezione, molte CPU prevedono un duplice modo di funzionamento (dual-mode):

- **kernel mode**: può eseguire istruzioni privilegiate. Il sistema operativo esegue in modo kernel a differenza dei processi utente che per eseguire istruzioni privilegiate si avvalgono di system call;
- **user mode**: programmi utente.

1.6 Strutture

Le componenti di un sistema operativo possono essere organizzate in modi differenti all'interno del sistema operativo:

- **struttura monolitica:** il sistema è costituito da un unico modulo che realizza le varie componenti. Questo assicura un basso costo di interazione tra le varie componenti, ma il sistema operativo è estremamente complesso e poco adattabile;
- **struttura modulare:** le varie componenti del sistema operativo vengono organizzate in moduli caratterizzate da interfacce ben definite. Ogni strato ha un insieme di funzionalità che vengono mostrate allo strato superiore mediante interfacce. Questa struttura permette l'astrazione e la modularità ma implementa una organizzazione gerarchica delle componenti (che non sempre è possibile) e una scarsa efficienza data dal costo di attraversamento dei vari livelli;
- **struttura a microkernel:** la struttura del nucleo (kernel) è ridotta al minimo (solo funzionalità base). Il resto del sistema operativo è rappresentato da processi utente. È un sistema molto affidabile e personalizzabile ma non molto efficiente a causa delle molte system call.

1.7 Macchine virtuali

Le macchine virtuali sono l'evoluzione dell'approccio a livelli. Virtualizzano sia l'hardware che il kernel del sistema operativo. Su una stessa macchina fisica danno l'illusione di elaboratori multipli, ciascuno in esecuzione sul suo processo privato e con le sue risorse private, messe a disposizione dal kernel del sistema operativo, che può essere diverso per processi diversi.

Una singola piattaforma hardware può essere condivisa da più sistemi operativi, ognuno dei quali è installato su una diversa macchina virtuale. Il disaccoppiamento è realizzato da un componente detto Virtual Machine Monitor di sistema che ha il compito di consentire la condivisione da parte di più macchine virtuali di un'unica piattaforma hardware. In particolare, una VMM di sistema integra le sue funzionalità in un sistema operativo leggero, costituendo un sistema posto direttamente sopra l'hardware.

Una VMM ospitata è invece una macchina virtuale che viene installata come applicazione sopra un sistema operativo esistente, opera nello spazio utente e accede all'hardware tramite system call.

2. Processi e thread

2.1 Processi

Un processo è l'unità di esecuzione all'interno del sistema operativo. Solitamente l'esecuzione è sequenziale, ma un sistema operativo multi-programmato consente l'esecuzione concorrente di più processi. Un processo durante la sua esistenza può trovarsi in vari stadi:

- **init**: stato di transizione durante il quale il processo viene caricato in memoria e vengono inizializzati i suoi dati;
- **ready**: il processo è pronto ad acquisire la CPU;
- **running**: il processo sta utilizzando la CPU;
- **waiting**: il processo è in sospenso in attesa di un evento;
- **terminated**: stato di transizione tra la terminazione e la deallocazione del processo. Il sistema operativo associa ad ogni processo una struttura dati detta Process Control Block. Il PCB contiene tutte le informazioni relative al processo.

2.2 Scheduling

Lo scheduling è l'attività mediante la quale il sistema operativo effettua delle scelte tra i processi, riguardo al loro caricamento in memoria centrale o l'assegnazione alla CPU. In generale il sistema operativo effettua tre diverse attività di scheduling:

- **scheduling a breve termine** (o di CPU): si occupa di selezionare il processo da assegnare alla CPU e di effettuare il context switch, ovvero la fase in cui l'uso della CPU viene commutato da un processo ad un altro;
- **scheduling a medio termine** (o swapping): gestisce il trasferimento temporaneo in memoria secondaria di processi (o parti di essi) in modo da consentire il caricamento di altri processi in memoria centrale;
- **scheduling a lungo termine**: seleziona i programmi da eseguire dalla memoria secondaria per caricarli in memoria centrale. È una componente importante dei sistemi batch e dei sistemi multi-programmati, ma non è presente nei sistemi time sharing in quanto è l'utente a stabilire i processi da eseguire.

2.3 Operazioni sui processi

Ogni sistema operativo prevede dei meccanismi per la gestione dei processi:

- creazione;
- interazione tra processi;
- terminazione.

Queste sono operazioni privilegiate e vengono eseguite in *Kernel mode*.

2.4 Thread (o processi leggeri)

Un thread (o processo leggero) è un'unità di esecuzione che condivide codice e dati con gli altri thread ad esso associati. L'insieme di thread che riferiscono allo stesso codice e dati è detto *task*. Un processo pesante equivale ad una task con un singolo thread.

I thread hanno come vantaggio la condivisione di memoria con altri thread dello stesso task e un costo minore di contest switch in quanto il PCB dei thread non contiene nessuna informazione relativa a codice e dati. I thread garantiscono una minore protezione in quanto thread di uno stesso task possono modificarsi i dati a vicenda.

2.5 Interazione tra processi

Due processi si possono dire:

- **processi interagenti**: l'esecuzione di uno è influenzata dall'esecuzione dell'altro e/o viceversa. In particolare, l'interazione può essere:
 - *cooperazione*: interazione prevedibile e desiderata per il raggiungimento di un fine comune;
 - *competizione*: interazione prevedibile ma indesiderata;
 - *interferenza*: interazione non prevista e non desiderata;
- **processi indipendenti**: l'esecuzione di uno non è influenzata dall'esecuzione dell'altro e/o viceversa.

L'interazione tra processi può essere diretta, specificando il processo che deve ricevere il messaggio (e conoscendo il processo che spedisce il messaggio nel caso del ricevente), o indiretta con modello a mailbox.

L'interazione tra processi può avvenire mediante:

- **memoria condivisa** (ambiente globale): il sistema operativo consente ai processi di condividere variabili e l'interazione avviene tramite esse;
- **scambio di messaggi** (ambiente locale): i processi non condividono variabili e interagiscono mediante la trasmissione e la ricezione di messaggi.

2.6 Processi in UNIX

In UNIX ogni processo ha il proprio spazio di indirizzamento completamente locale e non condiviso (ambiente locale). In UNIX ogni processo può entrare in due ulteriori stati, oltre quelli generali, che sono:

- **zombie**: il processo è terminato ma in attesa che il padre ne rilevi lo stato di terminazione;
- **swapped**: il processo o parte di esso è temporaneamente trasferito in memoria secondaria dallo scheduler a medio termine.

Il codice (text) dei processi è rientrante, quindi più processi possono condividere lo stesso codice. Codice e dati sono separati (modello a codice puro). Il sistema operativo gestisce una struttura dati globale in cui sono contenuti i puntatori ai codice (text) utilizzati, ed eventualmente condivisi da altri processi. Questa struttura è detta text table. L'elemento della text table è detto text structure e contiene un puntatore al codice e il numero di processi che lo condividono.

3. Scheduling della CPU

3.1 Generalità

Lo scheduler della CPU è quel componente del sistema operativo che si occupa di selezionare dalla coda dei processi *Ready* il prossimo processo al quale assegnare l'uso della CPU.

La gestione della coda dei processi *Ready* è realizzata mediante algoritmi di scheduling. In particolare, abbiamo due categorie di algoritmi di scheduling:

- senza prelazione: la CPU rimane allocata al processo *Running* finché esso non si sospende volontariamente o termina;
- con prelazione: il processo *Running* può essere prelazionato, cioè sospeso dal sistema operativo per assegnare la CPU ad un nuovo processo. I sistemi a divisione di tempo implementano sempre queste politiche di scheduling.

3.2 Criteri di scheduling

Per confrontare i vari algoritmi di scheduling vengono considerati alcuni indicatori di performance.

In generale deve essere massimizzato l'utilizzo della CPU e il numero di processi completati nell'unità di tempo; devono invece essere minimizzati il tempo di attesa di un processo nella *Ready queue*, il tempo tra l'inizio di un job e il suo completamento e l'intervallo tra la sottomissione di un job e la sua prima risposta.

3.3 Algoritmi di scheduling

- **FCFS** (*first come first served*): la coda dei processi è gestita in modalità FIFO. È un algoritmo senza prelazione ed è impossibile influire sull'ordine dei processi;
- **SJF** (*shortest job first*): per ogni processo *Ready* viene stimato il tempo di esecuzione e viene schedulato per primo il processo con il CPU burst stimato più breve. Può essere sia con e sia senza prelazione;
- **Round Robin**: tipicamente utilizzato nei sistemi time sharing, gestisce la coda come una FIFO circolare e ad ogni processo viene allocata la CPU per un intervallo di tempo costante detto *time slice*; dopo questo intervallo il processo viene reinserito in coda e la CPU passa al processo successivo;
- **con priorità**: ad ogni processo viene assegnata una priorità (definita dal sistema operativo o esternamente). La priorità di un processo può essere costante o variata dinamicamente. Questo algoritmo presenta il problema della *starvation* dei processi con priorità bassa. La soluzione è la modifica dinamica delle priorità, come ad esempio il decrescere della priorità al crescere del tempo assegnato alla CPU.

Spesso nei sistemi operativi reali vengono combinati diversi algoritmi di scheduling.

4. File System

4.1 Generalità

Il file system è quella componente del sistema operativo che fornisce i meccanismi di accesso e memorizzazione delle informazioni allocate in memoria secondaria. Realizza i concetti astratti di:

- **file**: unità logica di memorizzazione;
- **direttorio**: insieme di file;
- **partizione**: insieme di file associato ad un particolare dispositivo fisico.

L'organizzazione del file system è divisa su livelli, in particolare:

- **struttura logica**: presenta alle applicazioni una visione astratta delle informazioni memorizzate;
- **accesso**: definisce e realizza i meccanismi di accesso al contenuto del file, in particolare:
 - definisce il concetto di unità di trasferimento da/verso file: record logico;
 - definisce i metodi di accesso (sequenziale, diretto, ad indice);
 - realizza i meccanismi di protezione.
- **organizzazione fisica**: rappresentazione di file e directory sul disco e mapping dei record logici sui blocchi;
- **dispositivo virtuale**: vista astratta del dispositivo che appare come una sequenza di blocchi di dimensione costante.

4.2 File

Il file è un insieme di informazioni. Ogni file è individuato da almeno un nome simbolico ed è caratterizzato da un insieme di attributi:

- tipo;
- indirizzo;
- dimensione;
- data e ora di creazione e/o modifica;
- utente proprietario;
- protezione (diritti di accesso).

Tutti gli attributi sopra elencati sono contenuti all'interno di una struttura dati chiamata descrittore del file. Ogni descrittore deve essere memorizzato in modo persistente.

4.2.1 Operazioni sui file

Il compito del sistema operativo è quello di consentire l'accesso on-line ai file, ovvero di rendere immediatamente visibile le modifiche effettuate da un processo su un file agli altri processi. Le operazioni più comuni sono:

- **creazione**: allocazione di un file in memoria secondaria e inizializzazione dei suoi attributi;
- **lettura**: lettura di record logici dal file;
- **scrittura**: inserimento di nuovi record logici all'interno del file;
- **cancellazione**: eliminazione del file dal file system.

Per garantire una maggiore efficienza, il sistema operativo registra i file attualmente in memoria in una struttura detta tabella dei file aperti ed effettua lo spostamento dei file aperti in memoria centrale (*memory mapping*). Le operazioni necessarie sono:

- **apertura**: introduzione di un nuovo elemento nella tabella dei file aperti ed eventuale memory mapping;
- **chiusura**: salvataggio del file in memoria secondaria e rimozione dell'elemento corrispondente dalla tabella dei file aperti.

4.2.2 Struttura interna dei file

Ogni dispositivo di memorizzazione secondario viene diviso in blocchi di dimensione fissa (record fisici); essi sono le unità di trasferimento fisico nelle operazioni di I/O. L'utente vede il file come un insieme di record logici, ovvero le unità di trasferimento logico nelle operazioni di accesso ai file. I record logici hanno sempre dimensione variabile. È compito del sistema operativo stabilire una corrispondenza tra blocchi fisici e record logici. Di solito la dimensione dei blocchi è superiore a quella dei record logici: avviene quindi l'impacchettamento di più record logici all'interno dei blocchi.

4.2.3 Metodi di accesso ai file

L'accesso ai file può avvenire secondo tre modalità:

- **accesso sequenziale:** il file è una sequenza di record logici, quindi per accedere ad un particolare record è necessario accedere prima ai record che lo precedono;
- **accesso diretto:** il file è un insieme di record logici numerati; è molto utile quando si ha a che fare con file molto grandi per leggere/modificare poche informazioni;
- **accesso ad indice:** ad ogni file viene associata una struttura dati contenente l'indice delle informazioni contenute; per accedere ad un record logico si esegue una ricerca nell'indice.

4.3 Directory

La directory (o direttorio) è uno strumento per organizzare i file all'interno del sistema operativo. La struttura logica delle directory può variare a seconda del sistema operativo. In particolare, le realizzazioni logiche principali sono:

- **struttura ad un livello:** una sola directory per ogni file system (difficile separare file per una multi-utenza);
- **struttura a due livelli:** il primo livello contiene una directory per ogni utente del sistema, mentre il secondo livello (directory utenti) contiene i file dei rispettivi utenti;
- **struttura ad albero:** organizzazione gerarchica a N livelli. Ogni directory può contenere sia file che directory. Al primo livello si trova il direttorio radice;
- **struttura a grafo aciclico:** estende la struttura ad albero con la possibilità di inserire link differenti allo stesso file.

4.4 Organizzazione fisica del file system

Il sistema operativo si occupa anche della realizzazione del file system sui dispositivi di memorizzazione secondaria. In particolare, le principali tecniche di allocazione dei blocchi (insieme di record logici contigui) sono:

- **allocazione contigua:** ogni file è mappato su un insieme di blocchi fisicamente contigui dando la possibilità di accesso sequenziale e diretto;
 - vantaggi: basso costo di ricerca di un blocco e possibilità di accesso sequenziale e diretto;
 - svantaggi: difficile individuare lo spazio libero per l'allocazione di nuovi file. Frammentazione esterna: man mano che si riempie il disco, rimangono zone contigue sempre più piccole e a volte inutilizzate.
- **allocazione a lista:** i blocchi sui quali viene mappato ogni file sono organizzati a lista concatenata. I puntatori ai blocchi sono distribuiti sul disco;
 - vantaggi: non c'è frammentazione esterna e basso costo di allocazione;
 - svantaggi: necessità di dedicare spazio ai puntatori (maggiore spazio occupato), difficile rendere possibile l'accesso diretto e costo elevato per la ricerca di un blocco.
- **allocazione ad indice:** ad ogni file è associato un blocco (blocco indice) in cui sono contenuti tutti gli indirizzi dei blocchi su cui è allocato il file;
 - vantaggi: non c'è frammentazione esterna, basso costo di allocazione e possibilità di accesso diretto;

- svantaggi: possibile scarso utilizzo dei blocchi indice.

Esistono sistemi operativi che adottano più di un metodo di allocazione.

4.5 File system in UNIX

Il file system di UNIX è composto da tre tipologie distinte di file:

- file ordinari;
- directory;
- dispositivi fisici (file speciali contenuti nel direttorio */dev*).

I file vengono allocati con la tecnica di allocazione ad indice a più livelli di indirizzamento. Il file system è partizionato in quattro regioni:

- **boot block:** contiene le procedure di inizializzazione del sistema operativo;
- **super block:** fornisce i limiti delle quattro regioni, il puntatore alla lista dei blocchi liberi e il puntatore alla lista degli i-node liberi;
- **data blocks:** è l'area in cui memorizzare il file. Contiene i blocchi allocati e quelli liberi;
- **i-List:** contiene la lista di tutti i descrittori dei file (i-node) e i direttori e dispositivi presenti nel file system, accessibili attraverso l'indice i-number.

L'allocazione del file non è su blocchi fisicamente contigui; il suo i-node contiene i puntatori ai blocchi. Ogni file è accessibile secondo tre diverse modalità: scrittura, lettura ed esecuzione.

Il proprietario di un file può concedere o negare il permesso di accedere al file ad altri utenti. Esiste un utente privilegiato, detto utente root, che ha accesso ad ogni file del sistema. Ad ogni file, nel suo i-node, sono associati 12 bit di protezione. I primi 9 bit sono i permessi per l'utente proprietario, utenti nel gruppo e i restanti utenti; gli ultimi 3 bit di permessi sono per i file eseguibili.

Al processo che esegue un file eseguibile è associato dinamicamente un User ID e un Group ID. Anche i direttori sono rappresentati da file nel file system in UNIX.

Ogni file-direttorio contiene un insieme di record logici costituiti dal nome relativo di ogni file o sotto-direttorio contenuto ed il relativo i-number che lo identifica.

4.5.1 Accesso ai file in UNIX

UNIX implementa l'accesso sequenziale ai file. Un I/O pointer (puntatore al file) registra la posizione corrente nel file. Ad ogni processo è associata una tabella dei file aperti di processo di dimensione limitata. Ogni elemento della tabella rappresenta un file aperto dal processo ed è individuato da un indice intero detto file descriptor. I file descriptor rappresentano rispettivamente *stdin* (0), *stdout* (1) e *stderr* (2). La tabella dei file aperti del processo è allocata nella sua user structure.

Per realizzare l'accesso ai file, il sistema operativo utilizza due strutture dati globali, allocate nell'area dati del kernel:

- **tabella dei file attivi:** per ogni file aperto, contiene una copia del suo i-node;
- **tabella dei file aperti di sistema:** ha un elemento per ogni operazione di apertura (anche dello stesso file) relativo ai file aperti (e non ancora chiusi). Ogni elemento contiene I/O pointer (indica la posizione corrente all'interno del file) e un puntatore all'i-node del file nella tabella dei file attivi.

5. Gestione della memoria centrale

5.1 Generalità

A livello hardware ogni sistema ha disposizione un unico spazio di memoria accessibile direttamente da CPU e altri dispositivi. I compiti del sistema operativo nella gestione della memoria sono:

- **accesso alla memoria centrale:** vengono svolte load e store di dati e istruzioni. Gli indirizzi possono essere simbolici, logici e fisici. Ogni processo dispone di un proprio spazio di indirizzamento logico che viene allocato nella memoria fisica;
- **binding degli indirizzi:** viene associato ad ogni indirizzo logico o simbolico un indirizzo fisico. Può essere svolto staticamente (a tempo di compilazione) o a tempo di caricamento;
- **caricamento/collegamento dinamico:** l'obiettivo è l'ottimizzazione della memoria. Le funzioni possono essere usate da più processi simultaneamente e il sistema operativo deve gestire gli accessi allo spazio di altri processi;
- **gestire spazi di indirizzi logici di dimensione maggiore allo spazio fisico:** l'overlay è la soluzione al problema in cui la memoria disponibile non sia sufficiente ad accogliere codice e dati di un processo; esso mantiene in memoria codice e dati usati più di frequente e che sono necessari nella fase corrente. Codice e dati vengono quindi divisi in overlay che vengono caricati e scaricati dinamicamente.

5.2 Tecniche di allocazione in memoria centrale

Codice e dati dei processi possono essere allocati in memoria centrale in due modi:

- **allocazione contigua;**
- **allocazione non contigua.**

5.2.1 Allocazione contigua

Nel caso dell'allocazione contigua, abbiamo a sua volta due implementazioni principali:

- **allocazione contigua a partizione singola:** la parte di memoria per l'allocazione dei processi non è partizionata e solo un processo alla volta può essere allocato in memoria (no multi-programmazione);
- **allocazione contigua a partizioni multiple:** c'è multiprogrammazione e quindi necessità di proteggere codice e dati di ogni processo. Ad ogni processo viene associata una partizione.

Le partizioni possono essere:

- **fisse:** la dimensione di ogni partizione è fissata a priori e per ogni processo viene cercata una partizione libera di sufficiente dimensione. Ha come problema il crearsi di frammentazione interna dovuta al sottoutilizzo della partizione e un grado di multiprogrammazione limitato al numero di partizioni;
- **variabili:** le partizioni vengono allocate dinamicamente e dimensionate in base alla dimensione del processo. Quando un processo viene schedato dal sistema operativo, esso cercherà un'area di memoria abbastanza grande dove allocare dinamicamente la partizione associata. Viene eliminata la frammentazione interna in quanto le partizioni sono della dimensione richiesta, ma questa tecnica ha delle problematiche che risiedono nella scelta dell'area migliore in cui allocare il processo e nella frammentazione esterna.

5.2.2 Allocazione non contigua

Per eliminare la frammentazione esterna si passa all'allocazione non contigua, che si divide in:

- **paginazione:** lo spazio fisico viene considerato come un insieme di pagine, dette frame, di dimensione costante, sulle quali mappare porzioni dei processi da allocare. In questa modalità non si parla più di frammentazione esterna ed è possibile caricare in memoria un sottoinsieme delle pagine logiche di un processo. Il binding tra indirizzi logici e indirizzi fisici può essere realizzato mediante una tabella delle pagine, che ad ogni pagina logica associa la pagina fisica corrispondente.

Dato che la tabella può essere molto grande, e la traduzione della corrispondenza tra i due indirizzi dev'essere molto veloce, ci sono varie possibili soluzioni:

- su registri della CPU;
- in memoria centrale;
- memoria centrale e cache.

Quando lo spazio logico di indirizzamento di un processo è molto esteso si ha un elevato numero di pagine e una tabella delle pagine di grandi dimensioni. In questo caso si può utilizzare la paginazione su più livelli, che consiste nell'allocazione non contigua con paginazione anche della tabella delle pagine. Il tempo di accesso aumenta al crescere dei livelli di paginazione.

Per limitare l'uso della memoria alcuni sistemi operativi utilizzano un'unica struttura dati globale detta tabella delle pagine invertita, che ha un elemento per ogni frame. Ogni "pagina invertita" rappresenta un frame e contiene l'identificativo (il PID) del processo a cui è assegnato il frame. Il tempo di ricerca in questa soluzione è alto e la struttura rende difficile la rientranza, ovvero la condivisione di un frame tra processi diversi;

- **segmentazione:** si basa sul partizionamento dello spazio logico degli indirizzi di un processo in segmenti caratterizzati da nome e lunghezza. Ogni segmento viene allocato in memoria in modo contiguo e ad ognuno di essi il sistema operativo associa un intero attraverso il quale lo si può riferire.

È presente una tabella dei segmenti che ha una entry per ogni segmento che ne descrive l'allocazione in memoria fisica.

La segmentazione (più segmenti per processo) è l'evoluzione della tecnica di allocazione contigua a partizioni variabili (un segmento per processo). Il problema principale è la frammentazione esterna, che viene risolta con l'adozione di politiche adatte alla situazione.

Paginazione e segmentazione possono essere unite nella segmentazione paginata, in cui lo spazio logico è segmentato e ogni segmento è suddiviso in pagine. Con questo metodo si elimina la frammentazione esterna e non è necessario mantenere in memoria un intero segmento ma basta caricare solo le pagine necessarie. Sono presenti una tabella dei segmenti e una tabella delle pagine per ogni segmento.

5.3 Memoria Virtuale

La dimensione della memoria può rappresentare un vincolo importante riguardo alla dimensione dei processi e al grado di multiprogrammazione. La memoria virtuale è una tecnica che permette l'esecuzione di processi non completamente in memoria. Questa tecnica ha molti vantaggi:

- la dimensione dello spazio logico degli indirizzi è indipendente dall'estensione della memoria;
- il grado di multiprogrammazione è indipendente dalla dimensione della memoria fisica;
- il caricamento di un processo e il suo swapping hanno un costo ridotto;
- il programmatore non deve preoccuparsi dei vicoli relativi alla dimensione della memoria.

5.3.1 Paginazione su richiesta

Solitamente la memoria virtuale è realizzata mediante tecniche di paginazione su richiesta in cui tutte le pagine di ogni processo risiedono in memoria secondaria e durante l'esecuzione alcune di esse vengono trasferite all'occorrenza in memoria primaria. È presente un pager, ovvero una componente del sistema operativo che si occupa del trasferimento delle pagine. È detto *pager lazy* un pager che trasferisce in memoria centrale una sola pagina se ritenuta necessaria. Se si utilizza la paginazione su richiesta, l'esecuzione di un processo può richiedere swap-in del processo.

Una pagina dello spazio logico di un processo può essere allocata in memoria primaria o secondaria e la si distingue dai bit di validità presenti nella tabella delle pagine. La pagina può essere anche invalida, nel caso non esista lo spazio logico del processo. Nel caso di pagina invalida viene mandata un interrupt di *page fault*. In seguito al *page fault*, se è necessario caricare una pagina in memoria primaria, potrebbero non esserci frame liberi. In questo caso si utilizza la sovrallocazione, ovvero la sostituzione in memoria

primaria di una "pagina vittima" con la pagina richiesta. La "pagina vittima" viene quindi copiata in memoria secondaria, nel caso sia stata modificata. Si introduce quindi la necessità di introdurre efficaci ed efficienti algoritmi di sostituzione che siano in grado di sostituire quelle pagine la cui probabilità di accesso a breve termine è bassa. I principali sono:

- **LFU** (least frequently used): viene sostituita la pagina che è stata usata meno spesso;
- **FIFO** (first in, first out): viene sostituita la pagina che è da più tempo in memoria;
- **LRU** (least recently used): viene sostituita la pagina che è stata usata meno recentemente.

5.3.2 Pre-paginazione e working set

È possibile che il processo impieghi più tempo per la sua paginazione che per la sua esecuzione; in questo caso si parla di fenomeno di *thrashing*. Per contrastarlo si usano tecniche di gestione della memoria basate sulla pre-paginazione, ovvero nel prevedere il set di pagine, detto working set, di cui il processo da caricare avrà bisogno nella fase successiva di esecuzione. Un processo, in una sua fase di esecuzione, usa un sottoinsieme relativamente piccolo delle sue pagine logiche, che varia nel tempo.

Si dice località spaziale l'alta probabilità di accedere a locazioni vicine nello spazio a locazioni appena accedute, e località temporale l'alta probabilità di accedere a locazioni accedute di recente.

Nella tecnica della pre-paginazione con working set quando si carica un processo, si carica il suo working set iniziale, che verrà aggiornato dinamicamente in base al principio di località temporale.

5.4 Gestione della memoria in UNIX

UNIX utilizza la segmentazione paginata e la memoria virtuale tramite paginazione su richiesta senza working set. Inoltre, l'allocazione di ogni segmento non è contigua. La sostituzione delle pagine, tramite lo swapper di pagine, viene attivata dallo scheduler quando il numero totale di frame liberi è considerato insufficiente. Ogni frame ha dimensione costante.

6. Programmazione concorrente ad ambiente globale

6.1 Generalità

In una macchina ad ambiente globale le comunicazioni attraverso risorse condivise devono essere sincronizzate negli accessi. Esistono diverse tecniche di sincronizzazione gestite direttamente dal kernel del sistema operativo. Ogni applicazione può essere rappresentata come un insieme di due componenti:

- **processi** (componenti attivi);
- **risorse** (componenti passivi): sono oggetti fisici o logici che vengono utilizzati dal processo per il suo completamento. Le risorse sono raggruppate in classi di risorse: ogni classe identifica l'insieme delle operazioni che un processo può svolgere su tutte le risorse di quella classe. Le risorse possono essere private o globali a seconda che possano essere accedute e modificate da solo un processo o meno.

6.2 Deadlock

Un insieme di processi è in deadlock, o blocco critico, se ogni processo dell'insieme è in attesa di un evento che può essere causato solo da un altro processo dell'insieme: si ha quindi una situazione di stallo. Le condizioni necessarie per il deadlock sono quattro:

- mutua esclusione;
- possesso e attesa;
- impossibilità di prelazione;
- attesa circolare.

Se anche una delle quattro condizioni viene a mancare non c'è deadlock. Le situazioni di deadlock devono essere evitate o prevenendo il blocco critico o curandolo:

- **prevenzione**: può essere statica, nel caso in cui vengano posti vincoli ad almeno una delle quattro condizioni citate o dinamica, se il processo prima di cedere una risorsa controlla il possibile verificarsi del deadlock;
- **rilevazione/ripristino**: il sistema operativo rileva il deadlock e attua un algoritmo di ripristino.

6.3 Mutua esclusione

Il problema della mutua esclusione nasce quando più di un processo alla volta può avere accesso a risorse comuni. La regola di mutua esclusione impone che le operazioni con le quali i processi accedono alle risorse comuni non si sovrappongano nel tempo. In particolare, un'istruzione che opera su un dato è detta indivisibile (o atomica) se durante la sua esecuzione da parte di un processo, il dato non è accessibile da altri processi.

La sequenza di istruzioni con le quali un processo accede e modifica risorse comuni si dice sezione critica. La regola di mutua esclusione impone che una sola sezione critica di una classe possa essere in esecuzione in ogni istante. Lo schema generale di risoluzione è assicurare un prologo e un epilogo all'inizio e alla fine di ogni sezione critica. Il prologo in particolare è quell'insieme di istruzioni con il quale un processo richiede l'autorizzazione all'uso esclusivo di una risorsa. L'epilogo è l'insieme di istruzioni che vengono eseguite al completamento della sezione critica dove il processo dichiara libera la sezione critica.

Tra le tante soluzioni adottate l'algoritmo di Dekker è starvation-free: con il termine starvation si indica quando un processo attende un tempo infinito per l'acquisizione di una risorsa; per starvation-free si intende quindi un algoritmo che garantisce ad ogni processo di accedere alla sezione critica in un tempo finito.

7. Thread in Java

7.1 Generalità

Un thread condivide codice, dati e spazio di indirizzamento con gli altri thread associati; ha però stack e program counter privati. Ad ogni programma Java corrisponde l'esecuzione di un task (processo), contenente almeno un singolo thread, corrispondente al metodo `main()` sulla Java Virtual Machine. È comunque possibile creare dinamicamente thread che eseguono concorrentemente all'interno del programma. Essi sono implementabili o estendendo la classe `Thread` o implementando l'interfaccia `Runnable`. Il ciclo di vita di un thread può essere schematizzato in quattro fasi:

- **new thread**: subito dopo la creazione dell'oggetto;
- **runnable**: il thread è eseguibile ma potrebbe non essere in esecuzione;
- **not runnable**: stato nel quale finisce quando non può essere messo in esecuzione a causa di metodi quali `wait()`, `suspend()`, `sleep()` o perché in attesa di una operazione di I/O;
- **dead**: stato finale in cui giunge per "morte naturale" o perché è stato invocato il metodo `stop()`.

Le interazioni tra i thread avvengono mediante oggetti in comune. Si distinguono le interazioni:

- **competitive**: mutua esclusione;
- **cooperative**: semafori, variabili condizione.

7.2 Semafori

Il semaforo è uno strumento di sincronizzazione che permetti di risolvere qualunque problema di sincronizzazione tra thread nel modello ad ambiente globale. Viene implementato solitamente dal kernel del sistema operativo. Un semaforo è un dato astratto rappresentato da un intero non negativo a cui è possibile accedere solo tramite le operazioni `p()` e `v()`. Le due operazioni sono atomiche. Il valore del semaforo viene modificato da un processo alla volta.

Utilizzando i semafori, la decisione se un processo può proseguire l'esecuzione dipende solo dal semaforo e la scelta del processo da risvegliare avviene tramite un algoritmo FIFO. In problemi di sincronizzazione più complessi, invece, la decisione se un processo possa proseguire l'esecuzione dipende dal verificarsi di una condizione di sincronizzazione. Per ovviare a questi problemi, si utilizzano costrutti linguistici più di alto livello come i monitor.

7.3 Monitor

Il monitor è un costrutto sintattico che associa un insieme di operazioni dette entry ad una struttura dati comune a più thread, tale che le operazioni entry siano le sole operazioni permesse su quella struttura e siano mutuamente esclusive. Le variabili locali sono accessibili solo all'interno del monitor, così come le operazioni dichiarate non public.

Ad ogni monitor è associata una risorsa con lo scopo di controllare gli accessi in accordo a determinate politiche. L'accesso avviene mediante due livelli di sincronizzazione:

- solo un thread alla volta può accedere alle risorse del monitor tramite le operazioni entry;
- viene controllato l'ordine con cui i thread hanno accesso alle risorse del monitor in base ad una condizione di sincronizzazione che ne assicura l'ordinamento.

8. Gestione I/O

8.1 Compiti del gestore di I/O

- Nascondere al programmatore i dettagli delle interfacce hardware dei dispositivi;
- omogeneizzare la gestione di dispositivi diversi;
- gestire i malfunzionamenti che si possono verificare durante un trasferimento di dati;
- definire lo spazio dei nomi (naming) con cui vengono identificati i dispositivi, uniformandolo anche con il naming del file system;
- garantire la corretta sincronizzazione tra ogni processo applicativo che ha attivato un trasferimento dati e l'attività del dispositivo:
 - **gestione sincrona**: il processo si blocca fino al termine del trasferimento;
 - **gestione asincrona**: il processo attiva il trasferimento e procede nella sua esecuzione.

8.2 Architettura del sottosistema di I/O

La struttura del sottosistema di I/O è generalmente stratificata. I due livelli più importanti sono:

- il livello indipendente dai dispositivi, che implementa le funzioni di:
 - **naming**;
 - **buffering**: per ogni operazione di I/O il sistema operativo riserva un'area di memoria detta buffer per contenere i dati oggetti del trasferimento a causa della differenza di velocità tra processo e periferica, che provocherebbe un disaccoppiamento e della quantità dei dati da trasferire (che nei dispositivi a blocchi potrebbe essere inferiore alla dimensione del blocco);
 - **gestione dei malfunzionamenti**: prevede due approcci agli eventi anomali:
 - risoluzione del problema;
 - gestione parziale e propagazione al livello superiore (applicativo).

I tipi di eventi anomali che si possono generare sono:

- eventi propagati da un livello inferiore;
 - eventi generati a questo livello.
- **allocazione dei dispositivi ai processi applicativi**: deve gestire politiche di allocazione (nel caso di dispositivo condiviso da più processi: mutua esclusione).
- il livello dipendente dai dispositivi deve fornire i gestori dei dispositivi e offrire al livello superiore delle funzioni di accesso ai dispositivi. Contiene tutto il codice dipendente dal dispositivo. In particolare, la gestione dei dispositivi può essere implementata nei seguenti modi:
 - **a controllo di programma**: metodo non adatto per i sistemi multi-programmati, a causa dei cicli di attesa attiva;
 - **a interruzione**: per evitare l'attesa attiva ad ogni dispositivo viene assegnato un semaforo ed è il controller del dispositivo a dare il segnale di interruzione al sistema operativo quando è pronto per l'I/O;
 - **in DMA**: è un'alternativa che evita il sovraccarico della CPU rendendo possibile l'accesso "diretto" alla memoria tramite un controllore DMA. In questo modo viene bypassata la CPU e i suoi buffer.

8.3 Gestione e organizzazione fisica dei dischi

Un blocco fisico su un disco è descritto da tre parametri:

- **f**: numero della faccia;
- **t**: numero della traccia rispetto alla faccia;
- **s**: numero del settore all'interno della faccia.

Tutti i settori che compongono un disco (o un pacco di dischi) possono essere trattati come un array lineare di blocchi, indicando con M il numero di tracce per faccia e con N il numero di settori per traccia con l'indice i si

ha $i = (f \cdot M \cdot N) + (t \cdot N) + s$. Per ridurre il tempo medio di trasferimento di un settore si può intervenire in due modi:

- criteri con cui i dati sono memorizzati sul disco (metodi di allocazione dei file);
- criteri con cui servire le richieste di accesso (politiche di scheduling delle richieste).

Nella valutazione del tempo medio di attesa di un processo è necessario tenere conto anche il tempo durante quale il processo attende che la sua richiesta venga servita. Le richieste possono essere gestite secondo diverse politiche:

- **FCFS**: le richieste sono servite rispettando il tempo di arrivo;
- **SSTF**: seleziona la richiesta con il tempo di seek minimo a partire dalla posizione attuale della testina del disco. Può portare a situazioni di starvation delle richieste;
- **SCAN**: la testina si porta ad un'estremità del disco e si sposta verso l'altra servendo le richieste d'accesso man mano che vengono raggiunte le tracce indicate;
- **CSCAN**: è una variante di SCAN mirata a fornire un tempo di attesa medio dei processi più basso.